NC State University

Department of Electrical and Computer Engineering

ECE 463/563: Fall 2021 (Rotenberg)

Project #1: Cache Design, Memory Hierarchy Design

by

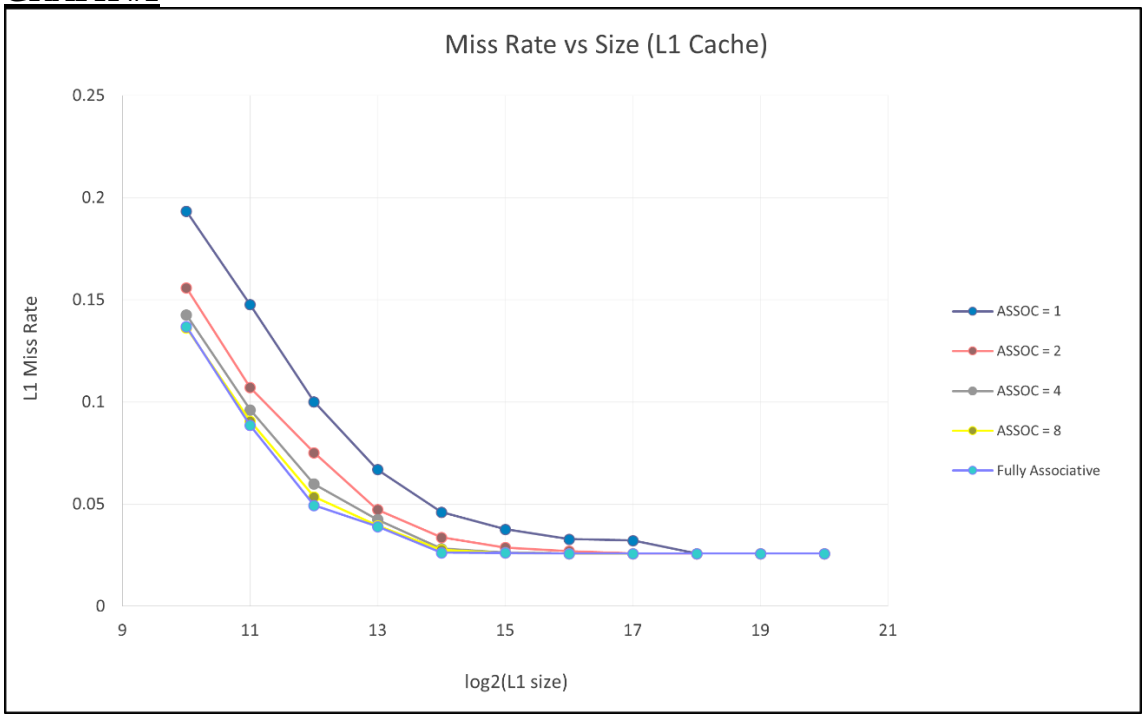
Cristian Hellmer

NCSU Honor Pledge: "I have neither given nor received unauthorized aid on this project."

Student's electronic signature: _____Cristian Hellmer_____
(sign by typing your name)

Course number: _____ECE 563_____

GRAPH #1



1. **Discuss trends in the graph. For a given associativity, how does increasing cache size affect miss rate? For a given cache size, what is the effect of increasing associativity?**

Increasing the L1 cache size for a given set-associativity decreases the miss rate. For a given cache size, increasing the associativity decreases the miss rate.

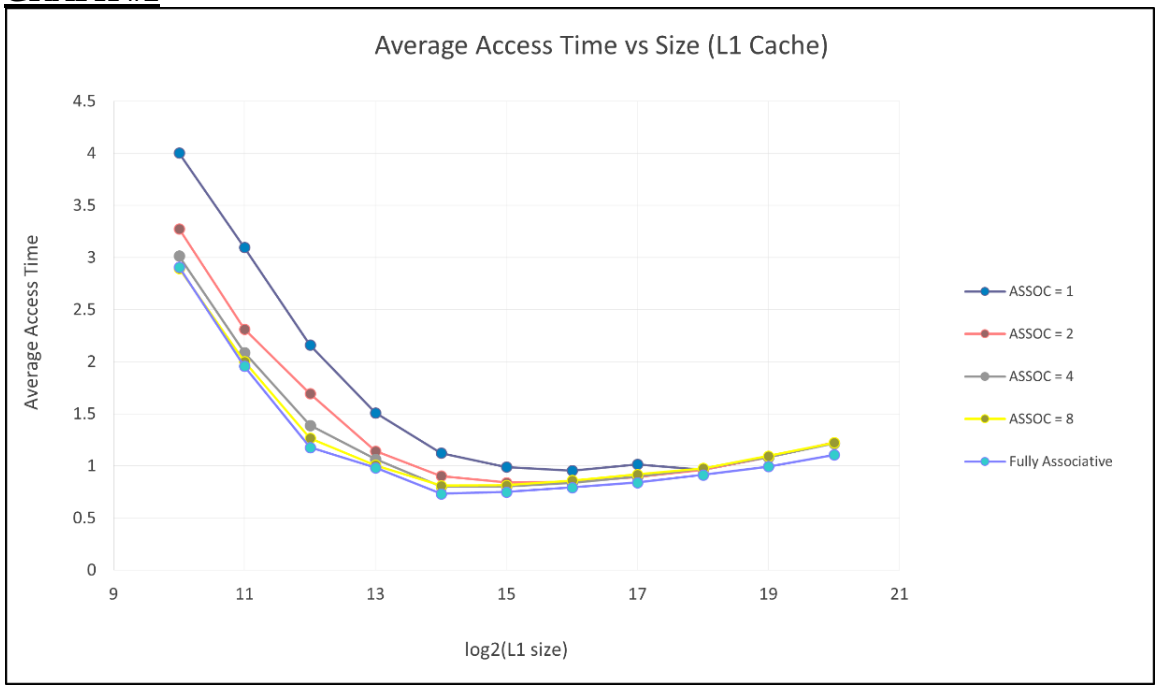
2. **Estimate the *compulsory miss rate* from the graph.**

Based on the graph, I estimate the *compulsory miss rate* to be 0.0258. This is because once you reach a certain size of cache for the larger associativity's, the miss rate does not decrease. This means that there probably isn't a capacity or conflict miss at these sizes, thus the entire miss rate is compulsory misses.

3. **For each associativity, estimate the *conflict miss rate* from the graph.**

Associativity	1	2	4	8	fully
Estimated *conflict miss rate*	0.0507	0.0258	0.0104	0.0041	0

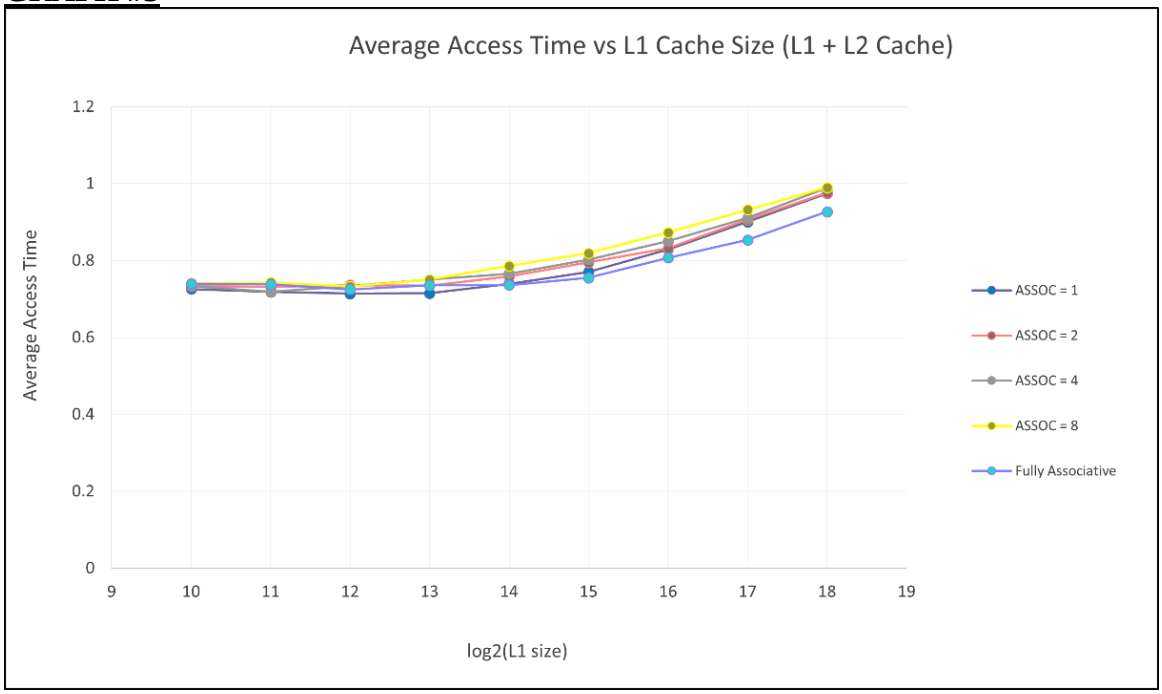
GRAPH #2



1. **For a memory hierarchy with only an L1 cache and BLOCKSIZE = 32, which configuration yields the best (*i.e.*, lowest) AAT?**

The lowest AAT for a memory hierarchy with only an L1 cache and BLOCKSIZE = 32 is achieved with a cache with SIZE = 16KB and ASSOC = 512. This cache is also fully associative. The AAT for this configuration is 0.734238ns.

GRAPH #3



1. **With the L2 cache added to the system, which L1 cache configurations result in AATs close to the best AAT observed in GRAPH #2 (*e.g.*, within 5%)?**

The L1 cache configurations resulting in AATs within 5% of the best AAT observed in GRAPH #2 include the following:

L1 Size	L1 Associativity
1024	1
2048	1
4096	1
8192	1
16384	1
32768	1
1024	2
2048	2
4096	2
8192	2
16384	2
1024	4
2048	4
4096	4
8192	4
16384	4
1024	8
2048	8
4096	8
8192	8
1024	32
2048	64
4096	128
8192	256
16384	512
32768	1024

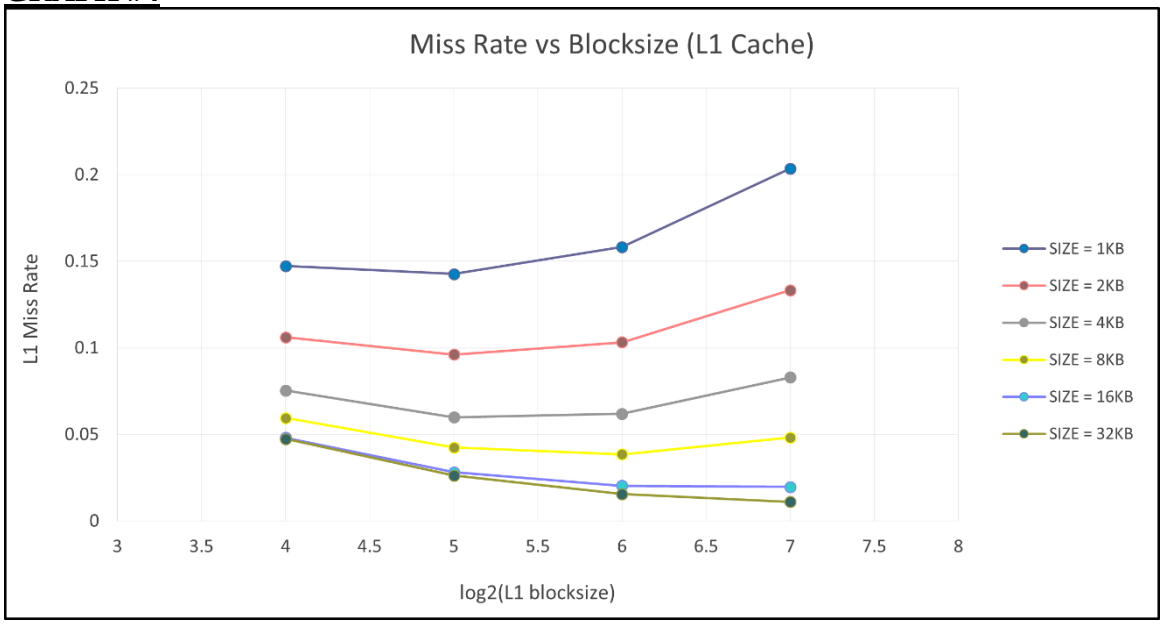
2. **With the L2 cache added to the system, which L1 cache configuration yields the best (*i.e.*, lowest) AAT? How much lower is this optimal AAT compared to the optimal AAT in GRAPH #2?**

With the L2 cache added to the system, the L1 cache configuration which yields the lowest AAT is achieved with a cache with SIZE = 4KB and ASSOC = 1. This optimal AAT is 0.020351ns lower than the optimal AAT from GRAPH #2.

3. **Compare the total area required for the optimal-AAT configurations with L2 cache (GRAPH #3) versus without L2 cache (GRAPH #2).**

	Without L2 Cache (GRAPH #2)	With L2 Cache (GRAPH #3)
Total Area	0.063446019 (mm*mm)	2.672838396 (mm*mm)

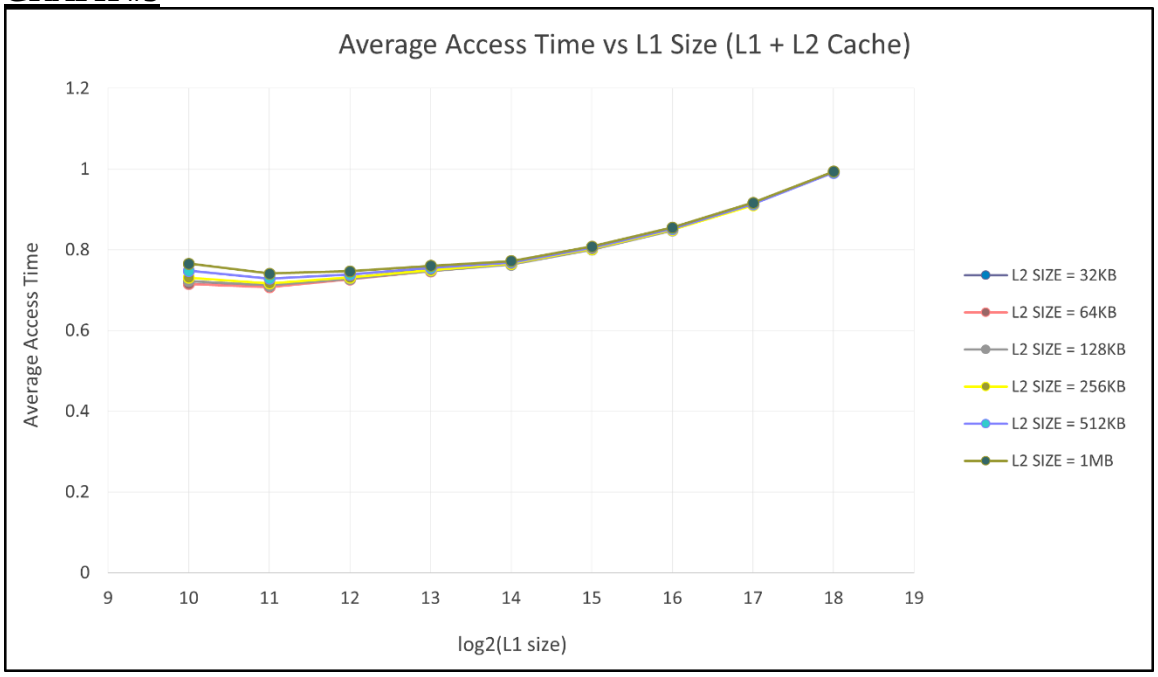
GRAPH #4



1. **Discuss trends in the graph. Do smaller caches prefer smaller or larger block sizes? Do larger caches prefer smaller or larger block sizes? Why? As block size is increased from 16 to 128, is the tradeoff between *exploiting more spatial locality* versus *increasing cache pollution* evident in the graph, and does the balance between these two factors shift with different cache sizes?**

Based on the graph, it seems that smaller caches prefer smaller block sizes while larger caches prefer larger block sizes. This is probably because smaller caches more quickly feel the effects of cache pollution which overwhelms the benefits from increasing special locality, reducing the number of blocks which can be used for hits on non-consecutive bytes. The balance shifts to larger block sizes when the cache size increases.

GRAPH #5



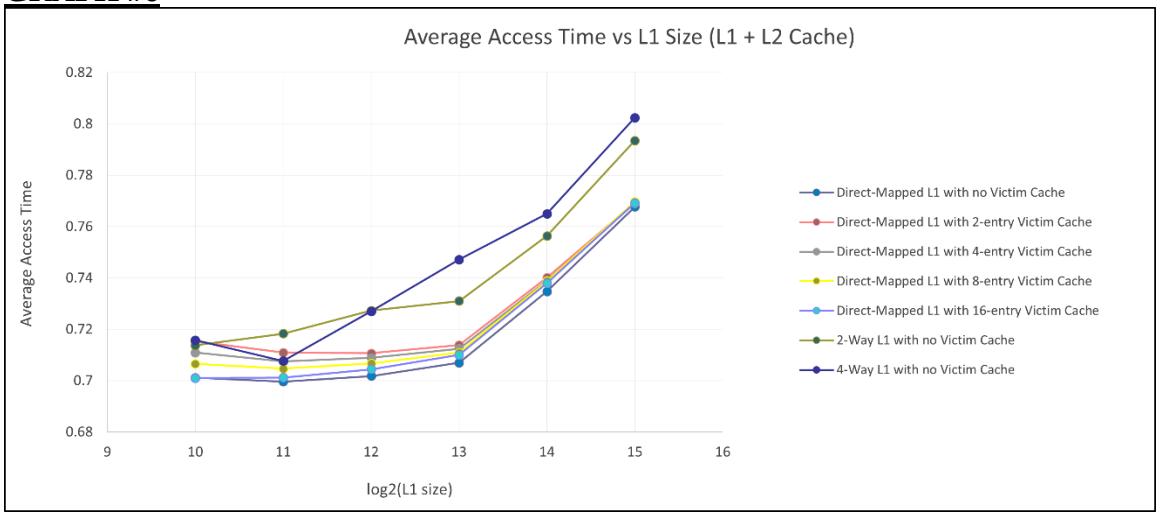
1. **Which memory hierarchy configuration yields the best (*i.e.*, lowest) AAT?**

The memory hierarchy configuration which yields the lowest AAT is with an L1 cache with SIZE = 2KB, ASSOC = 4, BLOCKSIZE = 32 and an L2 cache with SIZE = 64KB, ASSOC = 8, and BLOCKSIZE = 32. The AAT for this configuration is 0.707657833ns.

2. **Which memory hierarchy configuration has the smallest total area, that yields an AAT within 5% of the best AAT?**

The memory hierarchy configuration that has the smallest total area and yields an AAT within 5% of the best AAT is with an L1 cache with SIZE = 1KB, ASSOC = 4, BLOCKSIZE = 32 and an L2 cache with SIZE = 32KB, ASSOC = 8, and BLOCKSIZE = 32. This configuration has a total area of 0.257285583 (mm*mm). The AAT for this configuration is 0.7157522ns.

GRAPH #6



1. **Discuss trends in the graph. Does adding a Victim Cache to a direct-mapped L1 cache yield performance comparable to a 2-way set-associative L1 cache of the same size? …for which L1 cache sizes? …for how many Victim Cache entries?**

The direct-mapped L1 caches which included the Victim Cache outperformed the 2-way set associative L1 cache of the same size for almost every size.

2. **Which memory hierarchy configuration yields the best (*i.e.*, lowest) AAT?**

The memory hierarchy configuration which yields the lowest AAT is with an L1 cache with SIZE = 2KB, ASSOC = 1, BLOCKSIZE = 32, an L2 cache with SIZE = 64KB, ASSOC = 8, and BLOCKSIZE = 32, and no Victim Cache. The AAT for this configuration is 0.699615664ns.

3. **Which memory hierarchy configuration has the smallest total area, that yields an AAT within 5% of the best AAT?**

The memory hierarchy configuration that has the smallest total area and yields an AAT within 5% of the best AAT is with an L1 cache with SIZE = 1KB, ASSOC = 1, BLOCKSIZE = 32, an L2 cache with SIZE = 64KB, ASSOC = 8, and BLOCKSIZE = 32, and no Victim Cache. This configuration has a total area of 0.370616077 (mm*mm). The AAT for this configuration is 0.701216746ns.